

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan dalam bidang teknologi informasi telah maju sangat pesat, Indonesia sebagai negara berkembang turut merasakan dampak dari kemajuan ini. Salah satu dampak positifnya adalah penyebaran informasi yang dahulu membutuhkan waktu lama untuk tersebar, sekarang semakin cepat diterima seperti tidak dibatasi jarak, tetapi perkembangan ini juga memiliki dampak negatif, salah satunya adalah penyebaran berita hoaks. Menurut penelitian yang dikutip dalam [1], berita hoaks mempunyai peluang sampai 70% lebih tinggi untuk dibagikan dan tersebar lebih cepat di media sosial dibandingkan berita fakta. Kementerian Komunikasi dan Digital dalam siaran pers-nya menyatakan telah mengidentifikasi total sebanyak 11.759 berita hoaks dan informasi palsu sejak bulan Agustus 2018 sampai bulan Juni 2023, serta terdapat 1.923 kasus sepanjang tahun 2024 [2], [3]. Berita hoaks adalah ancaman dalam ranah keamanan siber, karena disinformasi berdampak pada integritas informasi [4]. Maka, penyebaran berita hoaks yang cepat ini jika tidak ditangani dengan baik akan menimbulkan permasalahan dalam keamanan informasi digital.

Menghadapi hal ini, diperlukan langkah penyelesaian untuk menjaga integritas data dengan cara melakukan deteksi berita hoaks sedini mungkin. Tetapi deteksi berita hoaks secara manual masih terbatas dan tidak mampu mengikuti kecepatan penyebaran informasi di media digital [5]. Maka diperlukan sistem otomatis deteksi berita hoaks yang bisa menjelaskan hasil keputusannya secara transparan yang dapat menjaga integritas dari suatu berita. Upaya ini juga sejalan dengan Schmitt, dkk yang berpendapat bahwa sistem deteksi berita hoaks harus transparan dan bisa menjelaskan keputusannya supaya mudah dipahami [6]. Hal ini juga relevan dengan arah kebijakan pemerintah Indonesia yang tercantum dalam Surat Edaran Menkominfo Nomor 9 Tahun 2023 tentang Pedoman Etika Kecerdasan Artifisial yang salah satunya berisi pedoman mengenai kredibilitas dan akuntabilitas penggunaan AI dimana informasi yang dihasilkan melalui Kecerdasan Artifisial harus dapat dipercaya dan dipertanggungjawabkan ketika

disebarkan kepada publik [7]. Selain itu, Kolaborasi Riset dan Inovasi Industri Kecerdasan Artifisial Indonesia (KORIKA) telah menerbitkan Strategi Nasional Kecerdasan Artifisial Indonesia 2020-2045 dan Pedoman Etika Kecerdasan Artifisial yang memberikan perhatian kepada transparansi keputusan sistem Kecerdasan Artifisial yang harus dapat dipahami manusia, dan memiliki akuntabilitas yang jelas dengan pengelolaan berbasis risiko [8], [9].

Pendekatan yang bisa dipakai adalah metode klasifikasi teks menggunakan *Natural Language Processing* (NLP) yang dapat digunakan untuk deteksi berita hoaks, salah satunya menggunakan *Bidirectional Encoder Representations from Transformers* (BERT) yaitu model *pre-trained Transformers* yang terbukti unggul dalam berbagai tugas NLP termasuk klasifikasi teks [10]. Sementara untuk

konteks bahasa Indonesia, model IndoBERT juga menunjukkan hal yang sama, seperti penelitian yang dilakukan oleh Tobing dkk. dan Ridho & Yulianti, menunjukkan bahwa model IndoBERT menghasilkan performa yang baik dalam mengklasifikasikan teks berita hoaks [11], [12]. Namun, sebagian besar penelitian terdahulu terkait IndoBERT dalam deteksi berita hoaks masih berfokus pada peningkatan performa klasifikasi seperti akurasi dan F1-Score, tanpa memberikan penjelasan mengenai bagaimana model menghasilkan keputusan prediksi. Padahal, dalam konteks deteksi berita hoaks, transparansi keputusan menjadi aspek penting karena hasil klasifikasi perlu dapat dipertanggungjawabkan dan diverifikasi. Ketidakmampuan model dalam menjelaskan alasan prediksi menyebabkan model bersifat *black box* dan menimbulkan permasalahan kepercayaan, akuntabilitas, serta kesulitan dalam proses validasi hasil prediksi.

Meskipun model IndoBERT menunjukkan performa yang baik, tetapi model ini bersifat *black box*, berarti proses pengambilan keputusannya tidak transparan dan sulit diinterpretasikan, sehingga menimbulkan permasalahan kepercayaan, akuntabilitas, dan kepatuhan terhadap standar etika dan hukum [13]. Dalam konteks sistem deteksi berita hoaks, tidak adanya transparansi dari proses pengambilan keputusan akan menyebabkan sulitnya memverifikasi benar atau tidaknya hasil prediksi model, karena berita hoaks memiliki ciri-ciri kebahasaan tertentu, seperti penggunaan makna tersirat yang memiliki maksud tersembunyi dan pilihan katanya cenderung manipulatif yang memprovokasi pembaca [14].

Maka dari itu, perlu untuk memahami kata seperti apa yang mempengaruhi hasil saat model memutuskan suatu teks itu hoaks atau bukan, tetapi IndoBERT tidak memberikan transparansi mengenai kata mana yang berkontribusi terhadap hasil prediksi. Keterbatasan pada model *black box* ini bisa diatasi oleh *Explainable Artificial Intelligence (XAI)*, dengan menggunakan *SHapley Additive exPlanations (SHAP)*. Metode SHAP dapat mengidentifikasi dan mengukur kontribusi fitur yang mempengaruhi hasil prediksi model [15]. Dengan menerapkan SHAP pada IndoBERT, penelitian ini memiliki tujuan untuk mempelajari kontribusi fitur yang mempengaruhi keputusan model pada konteks deteksi berita hoaks dengan penyajian yang informatif. Hal ini diharapkan mampu mengisi *gap* penelitian yang ada, dan dapat menghasilkan pemahaman yang lebih baik mengenai fitur apa saja yang berkontribusi terhadap hasil prediksi model IndoBERT, serta bisa mendukung upaya perlindungan keamanan informasi digital di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana menerapkan dan mengukur kinerja model IndoBERT dalam mendeteksi berita hoaks berbahasa Indonesia?
2. Bagaimana menerapkan metode SHAP pada model IndoBERT untuk mengidentifikasi fitur yang berpengaruh terhadap hasil prediksi?
3. Bagaimana menyajikan hasil interpretasi SHAP yang mengintegrasikan penjelasan naratif berbasis NLG Rule Based?

1.3 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan yang hendak dicapai, yaitu sebagai berikut:

1. Menerapkan dan mengukur kinerja model IndoBERT dalam mendeteksi berita hoaks berbahasa Indonesia.
2. Menerapkan metode SHAP pada model IndoBERT untuk mengidentifikasi fitur yang berpengaruh terhadap hasil prediksi.

3. Menyajikan hasil interpretasi SHAP dalam bentuk penjelasan naratif berbasis NLG Rule Based.

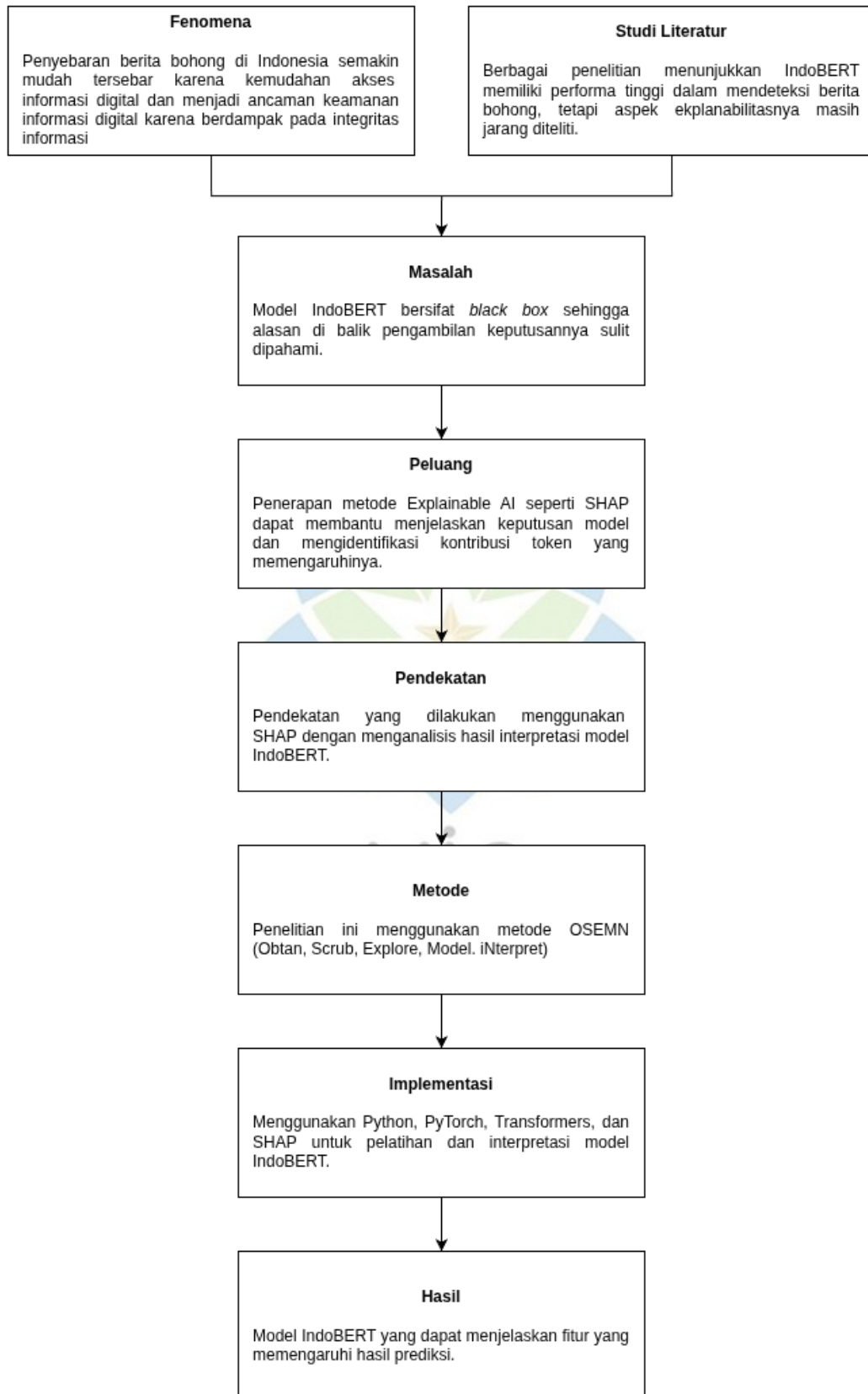
1.4 Batasan Masalah

Supaya penelitian yang dilakukan dapat terarah dengan baik, penelitian ini dibatasi pada hal-hal sebagai berikut:

1. Model yang digunakan adalah IndoBERT Base sebagai versi BERT untuk Bahasa Indonesia.
2. Metode XAI yang digunakan untuk menjelaskan keputusan model adalah SHAP.
3. Data yang digunakan adalah dari hasil *scraping* dari platform CNN, Kompas, Tempo, dan Turnbackhoax sebagai data primer dan dataset Indonesian *Fact and Hoax Political News* dari Kaggle sebagai data sekunder.
4. Pelabelan hasil klasifikasi dibatasi pada dua kategori yaitu hoaks dan fakta.
5. Evaluasi kontribusi fitur dilakukan hanya pada data test bukan keseluruhan dataset.
6. Input inferensi model berupa paragraf utuh, bukan kata tunggal.

1.5. Kerangka Pemikiran

Berdasarkan uraian sebelumnya, kerangka pemikiran dapat dilihat pada Gambar 1.1 yang bertujuan untuk menggambarkan alur penelitian.



Gambar 1.1 Alur Kerangka Pemikiran

1.6. Sistematika Penulisan

Sistematika penulisan ini disusun untuk memberikan gambaran umum mengenai penelitian yang dilakukan. Penelitian ini terdiri dari lima bab, adapun struktur dan penjelasan dari setiap bab adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas latar belakang permasalahan yang berkaitan dengan maraknya penyebaran berita hoaks serta urgensi penerapan *Explainable Artificial Intelligence* (XAI). Selain itu, bab ini juga memuat rumusan masalah, tujuan penelitian, batasan masalah yang ditetapkan, kerangka pemikiran yang digunakan, gambaran umum metode penelitian, serta sistematika penulisan skripsi.

BAB II KAJIAN LITERATUR

Bab tinjauan pustaka ini memuat pembahasan mengenai *state of the art* yang berisi rangkuman penelitian-penelitian terdahulu yang relevan sebagai acuan penelitian. Selain itu, bab ini juga menyajikan landasan teori yang menguraikan konsep, definisi, dan teori-teori yang digunakan sebagai dasar dalam pelaksanaan penelitian ini.

BAB III METODOLOGI

Bab ini menjelaskan metode penelitian yang digunakan dengan mengacu pada kerangka kerja OSEMN. Kerangka kerja tersebut terdiri atas tahapan *Obtain, Scrub, Explore, Model, dan Interpret* yang digunakan sebagai alur utama dalam proses pengolahan data serta pengembangan model pada penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan gambaran hasil penelitian yang diperoleh dari seluruh proses yang telah dilakukan. Selain itu, bab ini juga memuat analisis pada setiap tahapan OSEMN sebagai dasar dalam mengevaluasi hasil dan pembahasan penelitian.

BAB V SIMPULAN DAN SARAN

Bab ini memuat simpulan dari seluruh rangkaian penelitian yang telah dilakukan sebagai jawaban atas rumusan masalah yang telah ditetapkan. Selain itu, bab ini juga menyajikan saran-saran yang ditujukan untuk pengembangan penelitian selanjutnya guna mengatasi keterbatasan dan kekurangan yang ditemukan dalam penelitian ini.